BEBR

FACULTY WORKING PAPER NO. 1418

Nonparametric Kernel Estimation of Econometric Parameters

A. Ullah
H.D. Vinod

College of Commerce and Business Administration
Bureau of Economic and Business Research
University of Illinois, Urbana-Champaign

BEBR

FACULTY WORKING PAPER NO. 1418

College of Commerce and Business Administration

University of Illinois at Urbana-Champaign

December 1987

Nonparametric Kernel Estimation of Econometric Parameters

A. Ullah, Visiting Professor
Department of Economics

H.D. Vinod
Fordham Univesity

NONPARAMETRIC KERNEL ESTIMATION OF
ECONOMETRIC PARAMETERS

by

A. Ullah and H. D. Vinod*
University of Western Ontario, London, Ontario, and
Fordham University, Bronx, New York

ABSTRACT

We propose Nadaraya-Watson type nonparametric estimation of the conditional expectation of the dependent variable as a means of computing analytical partial derivatives (e.g., response coefficient, elasticity) with respect to appropriate variables. An illustrative example concerns the effect of age on earnings.

*Research support to A. Ullah from the NSERC is gratefully acknowledged. The authors thank Paul Rilstone and J. Racine for computational assistance and comments. They are also grateful to a referee for useful comments and suggestions.

1. INTRODUCTION AND THE MODEL

Let us consider an amorphous specification

$$y = R(x_1,\ldots,x_p) + u = E(y:x_1,\ldots,x_p) + u \qquad (1)$$

where y is an n x 1 vector of observations on the dependent variable, $x_1,\ldots,x_p$ are each n x 1 vector of observations on p regressors, u is an n x 1 vector of errors and the regression function R = R( ) is an unspecified expectation of y conditional on x_1 to x_p denoted by $E(y:x_1,\ldots,x_p)$. In the usual parametric econometrics, one considers R as a linear function of x_1 to x_p, $R = x_1\beta_1 + \ldots + x_p\beta_p$ where β_j, j = 1,...,p, is interpreted as the partial derivative (regression coefficient) of y with respect to x_j. Recent research has focused on achieving greater economic realism by using flexible functional forms, see e.g., Barnett (1984) and Elbadawi _et. al_., (1983).

In this paper we directly estimate the conditional expectation $R = E(y:x_1,\ldots,x_p)$ while its partial derivatives with respect to x_j for j = 1,...,p. The conditional expectation is estimated by the nonparametric Kernel method, and the estimation of the analytic partial derivatives appears to be new. This nonparametric approach to calculating partial derivatives has several advantages compared to the usual parametric approach. First, it does not require any _a priori_ assumption about the functional form. Second, x's are considered to be stochastic as is the case in the nonexperimental subject like economics. Third, it does not require any assumption about the data generating process (joint density of y, $x_1,\ldots,x_p$).

2. ESTIMATION OF PARTIAL DERIVATIVES

Let x_t, $t = 1,\ldots n$ be n independent and identically distributed random vectors generated from an unknown m-variate density function $f(x_1,\ldots,x_m)$. Consider K to be a class of all Borel-measurable real valued bounded functions k on the m-dimensional Euclidian space R^m such that

$$\int k(w)dw = 1, \qquad \int |k(w)|dw < \infty$$

$$\|w\|^m |k(w)| \to 0 \text{ as } \|w\| \to \infty \tag{2}$$

where $\|w\|$ is the usual Euclidian norm of w in R^m. Cacoullos (1966) estimated the joint density $f(x_j,\ j=1,\ldots,m)$ at point x_{jo} by

$$f_n(x_{10},\ldots,x_{mo}) = n^{-1}(\prod_1^m h_j^{-1}) \sum_1^n k(w_{1t},\ldots,w_{mt}) \tag{3}$$

where h_j is a sequence of positive numbers ("window widths") tending to zero as n tends to infinity and

$$w_{it} = (x_{jt}-x_{jo})/h_j. \tag{4}$$

Following Singh (1981) and Ullah and Singh (1985) we use, without the loss of generality, a joint kernel function k in (3) which is a product of m kernel functions $k(w_j)$, where each $k(w_j)$ satisfies (2). The particular choice of such kernel considered in Section 2.1 below is

$$k(w_{jt}) = (2\pi)^{-1/2}\exp[(-1/2)w_{jt}^2], \tag{5}$$

which is proportional to the normal density.

Let y be denoted by x_m with $m = p + 1$. Then the estimate of the marginal density $f(x_1,\ldots,x_p)$ at $x_{10},\ldots,x_{po}$, can be obtained by integrating out the variable $y = x_m$. Further, the estimate of the conditional mean $E(y:x_1,\ldots,x_p)$ or the regression function R in (1), at $x_{10},\ldots,x_{po}$, is

$$R_n(x_{10},\ldots,x_{po}) = R_n = \int x_{mo} f(x_{10},\ldots,x_{mo})/f(x_{10},\ldots,x_{po})dx_{mo}$$

$$= \sum_1^n y_t r_t;\ r_t = k(w_t)/\sum_1^n k(w_t) \tag{6}$$

where $k(w_t) = k(w_{1t},\ldots,w_{pt}) = \prod_1^p k(w_{jt})$. Note that $\sum_1^n r_t = 1$. This is the Nadaraya (1964) and Watson (1964) type estimator. The consistency, asymptotic properties of this estimator have been analyzed in Singh et. al., (1987) and Bierens (1987); also see Schuster (1972) and Rao ((1983), Ch. 4) for the special case $p = 1$.

We turn next to the nonparametric estimation of the partial derivatives of R with respect to x_j. The analytical expression for these derivatives follows from (6) as

$$\partial R_n/\partial x_{jo} = \sum_1^n y_t(K_{1t}-K_{2t}) \tag{7}$$

where

$$K_{1t} = k'(w_t)/\sum_1^n k(w_t) \tag{8}$$

$$K_{2t} = k(w_t) \sum_1^n k'(w_t)(\sum_1^n k(w_t))^{-2} \tag{9}$$

$$k'(w_t) = \partial k(w_t)/\partial x_{jo} = w_{jt} h_j^{-1} k(w_t) \tag{10}$$

where the second equality in (10) is true only for the normal kernel and k(w) is as given in (6). The asymptotic properties of the estimator in (7) follow directly from the asymptotic results in Singh et. al., (1987) and Schuster (1972), and they are not reproduced here for the sake of space. The interested reader may see the unpublished reports by Vinod and Ullah (1987), and Rilstone (1987) and Rilstone and Ullah (1987) which use numerical derivative instead of analytical expressions used in (7). We do, however, briefly outline the main results.

Using (1) in (6) we observe that

$$R_n = \sum_1^n R(x_{1t},\ldots,x_{pt})r_t + \sum_1^n u_t r_t = R + \sum_1^n u_t r_t + o_p(1) \qquad (11)$$

and, therefore, for large n

$$\frac{\partial R_n}{\partial x_{jo}} - \frac{\partial R}{\partial x_{jo}} \simeq \sum_1^n u_t(K_{1t}-K_{2t}) \qquad (12)$$

where the second equality in (11) follows by using the Taylor series expansion of $R(x_{1t},\ldots,x_{pt})$ around $x_{1o},\ldots,x_{po}$; $o_p(1)$ represents terms tending to zero in probability. Since the expectation of u conditional on x's is zero according to (1), the consistency of the estimator in (7) follows from (12). For the asymptotic distribution we note that, conditional on x's

$$Z = \left(\frac{\partial R_n}{\partial x_{jo}} - \frac{\partial R}{\partial x_{jo}}\right)/\Lambda^{1/2}(x_o) \sim N(0,1) \qquad (13)$$

where, from (12), $\Lambda(x_o) = \sum_1^n \sigma_u^2(x_t)(K_{1t}-K_{2t})^2$ is the asymptotic variance of $\partial R/\partial x_{jo}$ conditional on x's and $\sigma_u^2(x_t)$ is the conditional variance of u. Since the conditional distribution of Z in (13) is free from x's, the unconditional distribution of Z is also N(0,1). Note that Vinod and Ullah (1987) and Rilstone and Ullah (1987) have considered unconditional variance in the denominator of (13).

2.1 Numerical Example

As an illustration we analyze the response of earnings with respect to age. The partial derivative of the following semi-log nonparametric regression model provides the desired answer. Note that economists are specifically interested in the estimates of the partial derivatives, not the regression coefficients on a linear model, per se.

Suppose y_t is the logarithm of earnings and x_t is the age of the t^{th} individual. For simplicity in illustration, we assume schooling to be constant. Now the nonparametric specification of the model is

$$y_t = R(x_t) + u_t = E(y_t : x_t) + u_t \tag{14}$$

which is a special case of (1) for p = 1. Note that the estimate of R and its partial derivative with respect to x can be calculated by using (6) and (7), respectively. For the calculations the kernel used was as given in (5), and following Singh et. al., (1987) and Rao (1983, pp. 65-67) the window-width h taken was $sn^{-1/5}$ where $s^2 = \sum_1^n (x_t - \bar{x})^2/n$. Further, the conditional variance of the partial derivative $\Lambda(x_o)$ in (13) was obtained by using its consistent estimator $\sum_1^n \hat{\sigma}_u^2(x_t)(K_{1t} - K_{2t})^2$ where $\hat{\sigma}_u^2(x_t) = \sum_1^n \hat{u}_t^2 r_t$ is the weighted residual sum of squares (RSS) based on the nonparametric residual $\hat{u} = y - R_n$.

In the extensive labor econometrics literature the parametric specification of the model is

$$R = E(y_t : x_t) = \alpha + \beta x_t + \gamma x_t^2, \tag{15}$$

see Heckman and Polachek (1974) and Mincer (1974) among others. In this model the estimate of the partial derivative is

$$\hat{\beta} + 2\hat{\gamma} x_t \tag{16}$$

and its variance, conditional on x's, is given by

$$V(\hat{\beta}) + 4x_t^2 V(\hat{\gamma}) + 4x_t \operatorname{cov}(\hat{\beta},\hat{\gamma}), \tag{17}$$

where $\hat{\beta}$ and $\hat{\gamma}$ are the respective estimates of β and γ.

For the purpose of calculations, we considered Canadian data (1971 Canadian Census Public Use Tapes) on 205 individuals' ages and their earnings. These individuals were educated to grade 13. Below we present the parametric estimates based on ordinary least squares (OLS) and our nonparametric estimates of the partial derivatives.

	Partial Derivative	St. Error	RSS
Nonparametric Estimation	.0162	.002	63.54
Parametric OLS Estimation	.0189	.003	63.60

The OLS estimation of the parametric model (15) is

$$\underset{(.518)}{y = 10.041} + \underset{(.027)}{.173x} - \underset{(.003)}{.002x^2} \tag{18}$$

where the numbers in parentheses are standard errors.

The nonparametric and parametric partial derivative estimates given above are the means of the respective partial derivative estimates at various ages. These estimates and their corresponding RSS are quite similar, although the nonparametric standard error is smaller. This is not surprising since for the data under consideration the quadratic parametric specification in (15) is fairly close to the true nonparametric specification, see Ullah (1985). The nonparametric specification in Ullah (1985), however, indicates a "dip" around the mean age with the result that when the nonparametric partial derivative was calculated at the mean age of 39, it gave the value -.008, while the parametric estimate remained the same as expected.

Another point to be noted is that if an investigator incorrectly specifies the parametric model (15) as $R = \alpha + \beta x_t + u_t$, then the partial derivative estimate will be approximately .011 which will be away from the OLS value above of .0189. Since this bias problem due to the misspecification of the functional form does not arise in the nonparametric approach, applied econometricians may find nonparametric estimation attractive in a variety of other problems. Of course, certain modern "tests" on specification may reveal the inadequacy of the OLS, but may not reveal a practical alternative afforded by our nonparametric approach.

The potential users should, however, be well aware of the following limitations of the nonparametric approach. We know very little about the reliability of tests obtained by this approach in finite samples. The standard errors could be imprecise when the number of regressors is large and or sample is small. Not much is known about the selection of window width in small samples. For further details about these limitations and the areas of future research, see Ullah (1987).

REFERENCES

Barnett, W. A. (1984), "The Minflex-Laurent Translog Flexible Functional Form," Discussion Paper 84-7, Center for Economic Research, University of Texas at Austin.

Bierens, H. J. (1987), "Kernel Estimators of Regression Function," forthcoming in Advances in Econometrics.

Cacoullos, T. (1966), "Estimates of a Multivariate Density," Annals of the Institute of Statistical Mathematics 18, 176-189.

Elbadawi, I., A. R. Gallant and G. Souza, (1983), "An Elasticity can be Estimated Consistently Without A Priori Knowledge of Functional Form," Econometrica 51, 1731-1751.

Heckman, J. and S. Polachek (1974), "Empirical Evidence on the Functional Form of the Earning Schooling Relationship," Journal of the American Statistical Association 69, 350-354.

Mincer, J. (1974), Schooling, Experience and Earnings, NBER (New York: Columbia Univesrity Press).

Nadaraya, E. A. (1964), "On Estimating Regression," Theory of Probability and Applications 9, 141-142.

Rao, B. L. S. P. (1983), Nonparametric Functional Estimation, New York: Academic Press.

Rilstone, P. (1987), Nonparametric Partial Derivative Estimation, Unpublished Ph.D. Dissertation, Department of Economics, University of Western Ontario.

Rilstone, P. and A. Ullah (1987), Nonparametric Estimation of Response Coefficients, Mimeo, Department of Economics, University of Western Ontario.

Schuseter, E. F. (1972), "Joint Asymptotic Distribution of the Estimated Regression Function at a Finite Number of Distinct Points," Annals of Mathematical Statistics 43, 84-88.

Singh, R. S. (1981), "Speed of Convergence in Nonparametric Estimation of a Multivariate mu-Density and Its Mixed Partial Derivatives," Journal of Statistical Planning and Inference 5, 287-298.

Singh, R. S., A. Ullah and R. A. L. Carter (1987), "Nonparametric Inference in Econometrics: New Applications, in I. MacNeill and G. Umphrey (eds.) Time Series and Econometric Model, D. Reidell, Holland.

Ullah, A. (1985), "Specification Analysis of Econometric Models," Journal of Quantitative Economics 1, 187-209.

Ullah, H. and R. S. Singh (1985), "The Estimation of Probability Density Functions and Its Applications in Econometrics," Technical Report 6, The Centre for Decision Sciences and Econometrics, University of Western Ontario.

Ullah, A. (1987), "Nonparametric Estimation of Econometric Functionals," Technical Report 17, University of Western Ontario.

Vinod, H. D. and A. Ullah (1987), "Flexible Production Function Estimation by Nonparametric Kernel Estimators," Technical Report 16, University of Western Ontario.

Watson, G. S. (1964), "Smooth Regression Analysis," Sankhya, Series A 26, 359-372.